

WHAT IS CLAIMED IS:

1. A network camera system, comprising:

a network terminal;

at least one network camera connected to the network terminal via a network; and

wherein the network camera comprises:

a camera unit;

a microphone;

a program transmitter, which transmits an applet or a plugin to the network terminal;

the network camera transmits a web page attached with an image data and/or an audio data, to the network terminal; and

wherein the network terminal, which operable by the applet or the plugin to reproduce voice based on the audio data which associated with the image data.

2. The network camera system according to claim 1, wherein the applet or the plugin reproduces only a voice based on the audio data with regard to a uppermost window in a plurality of image display window displayed in the network terminal.

3. The network camera system according to claim 1, wherein the applet or the plugin indicates a display sequence input button, which operable to input a window display sequence on an image displaying window screen displayed in the network terminal, and

the audio data is adjusted and reproduced in accordance with the window display sequence inputted by the display sequence input button.

4. The network camera system according to claim 1, wherein the applet or the plugin indicates a audio reproduction start button and a audio reproduction stop button on an image displaying window screen displayed in the network terminal, and output and stop of the audio data are respectively selected in accordance with inputs through the audio reproduction start button and the audio reproduction stop button.

5. The network camera system according to claim 1, wherein the applet or the plugin computes a distance between a center position of each image displaying window displayed in the network terminal and a center position of a display device of the network terminal, and the audio data is adjusted and reproduced in accordance with the computed distances in a case where a plurality of windows are displayed.

6. A network camera connected to a network terminal, the network camera comprising:

- a camera unit, which photographs an image data;
- a microphone, which collects a audio data; and
- a program transmitter, which transmits an applet or a plugin for reproducing voice based on the audio data which associated

with the image data in the network terminal, to the network terminal.

7. The network camera according to claim 6, further comprising a loudspeaker, which reproduces voice based on the audio data transmitted from the network terminal.

8. A network terminal connected to at least one network camera, comprising:

browser, which is capable of receiving a Web page from the network camera, in connection with a network;

display controller, which is capable of window-displaying an image data;

audio controller, which reproduces a voice based on a audio data; and

audio function extension unit, which extends a function of the browser and reproduces the voice based on the audio data which associated with the image data, in a case where the web page has been received.

9. A audio reproduction method comprising the steps of:

transmitting a Web page attached with a image data photographed by each network camera and a audio data, to a network terminal;

attaching the web page with an applet or a plugin;

reproducing a voice based on the audio data which associated

with the image data in the network terminal, by the applet or the plugin.

10. The audio reproduction method according to claim 9, further comprising the steps of:

reproducing only a voice based on the audio data with regard to uppermost window in a plurality of image displaying windows displayed in the network terminal, by the applet or the plugin.

11. The audio reproduction method according to claim 9, further comprising the steps of:

indicating an display sequence input button, which is capable of inputting a window display sequence, on an image data displaying on an image displaying window screen displayed in the network terminal;

weighting each audio data in accordance with the window display sequence inputted through the display sequence input button; and

adjusting and reproducing the each audio data in accordance with the weight,

wherein the indicating and weighting and the adjusting and the reproducing is carried by the applet or the plugin.

12. The audio reproduction method according to claim 9, further comprising the steps of:

computing a distance between a center position of each image

displaying window displayed in the network terminal and a center position of a display device of the network terminal;

weighting the voices in accordance with the computed distance in a case where a plurality of windows are displayed; and

adjusting and reproducing the audio data in accordance with the weights,

wherein the computing and weighting and the adjusting and the reproducing is carried out by the applet or the plugin.

13. The audio reproduction method according to claim 9, further comprising the steps of:

indicating a audio reproduction start button and a audio reproducing stop button on an image displaying window screen displayed in the network terminal; and

selecting output and stop of the audio data in accordance with inputs through the audio reproduction start button and the audio reproducing stop button,

wherein the indicating and selecting is carried by the applet or the plugin.

14. A program in the for audio reproduction, comprising:

an interface, which has functions of permitting a computer to access a plurality of imaging servers in compliance with requests from a browser, and receives audio data from the

respective imaging servers;

nesting acquisition section, which acquires display sequence information of individual Web pages transmitted from the plurality of imaging servers; and

a audio selector, which selects and reproduces the audio data in accordance with the display sequence information of the plurality of Web pages acquired by the nesting acquirer.

15. The program according to claim 14,

wherein the audio selector selects only the audio data which corresponds to the imaging server of the Web page whose display sequence information specifies a uppermost position.

16. The program according to claim 14,

wherein instead of the selection of the audio data, the selector weights the respective audio data received from the plurality of imaging servers, on the basis of the display sequence information, and it reproduces the audio data in accordance with the weights.